

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: OPPER, Martin  
BOSSLET, Klaus  
CZECH, Joerg
- (ii) TITLE OF INVENTION: CYTOPLASMIC EXPRESSION OF ANTIBODIES,  
ANTIBODY FRAGMENTS AND ANTIBODY FRAGMENT FUSION MOLECULES  
IN E. COLI
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Foley & Lardner
  - (B) STREET: 3000 K Street, N.W., Suite 500
  - (C) CITY: Washington
  - (D) STATE: D.C.
  - (E) COUNTRY: USA
  - (F) ZIP: 20007-5109
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/630,820
  - (B) FILING DATE: 10-APR-1996
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: DE 19513676.4
  - (B) FILING DATE: 11-APR-1995
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: GRANADOS, Patricia D.
  - (B) REGISTRATION NUMBER: 33,683
  - (C) REFERENCE/DOCKET NUMBER: 18748/306
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (202)672-5300
  - (B) TELEFAX: (202)672-5399
  - (C) TELEX: 904136

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Enterobacteriaceae: Escherichia coli
- (vii) IMMEDIATE SOURCE:

(B) CLONE: KS + E.c.-Beta-Gluc

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AAGCTTTCAT TGTTCGCTC CCTGCTGCGG

30

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Enterobacteriaceae: Escherichia coli

(vii) IMMEDIATE SOURCE:

(B) CLONE: KS + E.c.-Beta-Gluc

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TCTAGACCAT GGTACGTCCT GTACAAACCC CA

32

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:

(B) CLONE: pAB Stop c-DNA HC/hu-Beta-Gluc

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAATTCCATG GAACCAGAAC CAGAACCGAG CTCAACTCT

39

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: pAB Stop c-DNA HC/hu-Beta-Gluc

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TCTAGATAAC GAGGGCAAAA AATGGAGGTC CAACTGCAGG AGAGC

45

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3169 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Enterobacteriaceae: Escherichia coli
  - (B) STRAIN: pRAJ210
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: pTrc99 dicistr. Fab/E.c.-Beta-Gluc
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION:3..641
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION:666..3162

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CC ATG GAC ATC CAG ATG ACC CAG AGC CCA AGC AGC CTG AGC GCC AGC	47
Met Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser	
1 5 10 15	
GTG GGT GAC AGA GTG ACC ATC ACC TGT AGT ACC AGC TCG AGT GTA AGT	95
Val Gly Asp Arg Val Thr Ile Thr Cys Ser Thr Ser Ser Ser Val Ser	
20 25 30	
TAC ATG CAC TGG TAC CAG CAG AAG CCA GGT AAG GCT CCA AAG CTG CTG	143
Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu	
35 40 45	
ATC TAC AGC ACA TCC AAC CTG GCT TCT GGT GTG CCA AGC AGA TTC AGC	191
Ile Tyr Ser Thr Ser Asn Leu Ala Ser Gly Val Pro Ser Arg Phe Ser	
50 55 60	

GGT AGC GGT AGC GGT ACC GAC TTC ACC TTC ACC ATC AGC AGC CTC CAG Gly Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln 65 70 75	239
CCA GAG GAC ATC GCC ACC TAC TAC TGC CAT CAG TGG AGT AGT TAT CCC Pro Glu Asp Ile Ala Thr Tyr Tyr Cys His Gln Trp Ser Ser Tyr Pro 80 85 90 95	287
ACG TTC GGC CAA GGG ACC AAG GTG GAA ATC AAA CGT ACT GTG GCT GCA Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala 100 105 110	335
CCA TCT GTC TTC ATC TTC CCG CCA TCT GAT GAG CAG TTG AAA TCT GGA Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly 115 120 125	383
ACT GCC TCT GTT GTG TGC CTG CTG AAT AAC TTC TAT CCC AGA GAG GCC Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala 130 135 140	431
AAA GTA CAG TGG AAG GTG GAT AAC GCC CTC CAA TCG GGT AAC TCC CAG Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln 145 150 155	479
GAG AGT GTC ACA GAG CAG GAC AGC AAG GAC AGC ACC TAC AGC CTC AGC Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser 160 165 170 175	527
AGC ACC CTG ACG CTG AGC AAA GCA GAC TAC GAG AAA CAC AAA GTC TAC Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr 180 185 190	575
GCC TGC GAA GTC ACC CAT CAG GGC CTG AGC TCG CCC GTC ACA AAG AGC Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser 195 200 205	623
TTC AAC AGG GGA GAG TGT TAGTCTAGAT AACGAGGGCA AAAA ATG GAG GTC Phe Asn Arg Gly Glu Cys Met Glu Val 210 1	674
CAA CTG CAG GAG AGC GGT CCA GGT CTT GTG AGA CCT AGC CAG ACC CTG Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg Pro Ser Gln Thr Leu 5 10 15	722
AGC CTG ACC TGC ACC GTG TCT GGC TTC ACC ATC AGC AGT GGT TAT AGC Ser Leu Thr Cys Thr Val Ser Gly Phe Thr Ile Ser Ser Gly Tyr Ser 20 25 30 35	770
TGG CAC TGG GTG AGA CAG CCA CCT GGA CGA GGT CTT GAG TGG ATT GGA Trp His Trp Val Arg Gln Pro Pro Gly Arg Gly Leu Glu Trp Ile Gly 40 45 50	818
TAC ATA CAG TAC AGT GGT ATC ACT AAC TAC AAC CCC TCT CTC AAA AGT Tyr Ile Gln Tyr Ser Gly Ile Thr Asn Tyr Asn Pro Ser Leu Lys Ser 55 60 65	866
AGA GTG ACA ATG CTG GTA GAC ACC AGC AAG AAC CAG TTC AGC CTG AGA Arg Val Thr Met Leu Val Asp Thr Ser Lys Asn Gln Phe Ser Leu Arg 70 75 80	914
CTC AGC AGC GTG ACA GCC GCC GAC ACC GCG GTC TAT TAT TGT GCA AGA Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg 85 90 95	962
GAA GAC TAT GAT TAC CAC TGG TAC TTC GAT GTC TGG GGT CAA GGC AGC Glu Asp Tyr Asp Tyr His Trp Tyr Phe Asp Val Trp Gly Gln Gly Ser 100 105 110 115	1010

CTC	GTC	ACA	GTC	ACA	GTC	TCC	TCA	GCT	TCC	ACC	AAG	GGC	CCA	TCG	GTC	1058
Leu	Val	Thr	Val	Thr	Val	Ser	Ser	Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	
			120						125					130		
TTC	CCC	CTG	GCG	CCC	TGC	TCC	AGG	AGC	ACC	TCT	GGG	GGC	ACA	GCG	GCC	1106
Phe	Pro	Leu	Ala	Pro	Cys	Ser	Arg	Ser	Thr	Ser	Gly	Gly	Thr	Ala	Ala	
			135					140					145			
CTG	GGC	TGC	CTG	GTC	AAG	GAC	TAC	TTC	CCC	GAA	CCG	GTG	ACG	GTG	TCG	1154
Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	Glu	Pro	Val	Thr	Val	Ser	
		150					155					160				
TGG	AAC	TCA	GGC	GCC	CTG	ACC	AGC	GGC	GTG	CAC	ACC	TTC	CCG	GCT	GTC	1202
Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	His	Thr	Phe	Pro	Ala	Val	
	165					170					175					
CTA	CAG	TCC	TCA	GGA	CTC	TAC	TCC	CTC	AGC	AGC	GTG	GTG	ACC	GTG	CCC	1250
Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val	Pro	
180					185					190					195	
TCC	AGC	AGC	TTG	GGC	ACC	CAG	ACC	TAC	ACC	TGC	AAC	GTG	AAT	CAC	AAG	1298
Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Thr	Cys	Asn	Val	Asn	His	Lys	
			200						205					210		
CCC	AGC	AAC	ACC	AAG	GTG	GAC	AAG	AGA	GTT	GAG	CTC	GGT	TCT	GGT	TCT	1346
Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Arg	Val	Glu	Leu	Gly	Ser	Gly	Ser	
			215					220					225			
GGT	TCC	ATG	GTA	CGT	CCT	GTA	GAA	ACC	CCA	ACC	CGT	GAA	ATC	AAA	AAA	1394
Gly	Ser	Met	Val	Arg	Pro	Val	Glu	Thr	Pro	Thr	Arg	Glu	Ile	Lys	Lys	
		230					235					240				
CTC	GAC	GGC	CTG	TGG	GCA	TTC	AGT	CTG	GAT	CGC	GAA	AAC	TGT	GGA	ATT	1442
Leu	Asp	Gly	Leu	Trp	Ala	Phe	Ser	Leu	Asp	Arg	Glu	Asn	Cys	Gly	Ile	
	245					250					255					
GAT	CAG	CGT	TGG	TGG	GAA	AGC	GCG	TTA	CAA	GAA	AGC	CGG	GCA	ATT	GCT	1490
Asp	Gln	Arg	Trp	Trp	Glu	Ser	Ala	Leu	Gln	Glu	Ser	Arg	Ala	Ile	Ala	
260					265					270					275	
GTG	CCA	GGC	AGT	TTT	AAC	GAT	CAG	TTC	GCC	GAT	CCA	GAT	ATT	CGT	AAT	1538
Val	Pro	Gly	Ser	Phe	Asn	Asp	Gln	Phe	Ala	Asp	Ala	Asp	Ile	Arg	Asn	
			280					285						290		
TAT	GCG	GGC	AAC	GTC	TGG	TAT	CAG	CGC	GAA	GTC	TTT	ATA	CCG	AAA	GGT	1586
Tyr	Ala	Gly	Asn	Val	Trp	Tyr	Gln	Arg	Glu	Val	Phe	Ile	Pro	Lys	Gly	
			295					300					305			
TGG	GCA	GGC	CAG	CGT	ATC	GTG	CTG	CGT	TTC	GAT	GCG	GTC	ACT	CAT	TAC	1634
Trp	Ala	Gly	Gln	Arg	Ile	Val	Leu	Arg	Phe	Asp	Ala	Val	Thr	His	Tyr	
		310					315					320				
GGC	AAA	GTG	TGG	GTC	AAT	AAT	CAG	GAA	GTG	ATG	GAG	CAT	CAG	GGC	GGC	1682
Gly	Lys	Val	Trp	Val	Asn	Asn	Gln	Glu	Val	Met	Glu	His	Gln	Gly	Gly	
	325					330					335					
TAT	ACG	CCA	TTT	GAA	GCC	GAT	GTC	ACG	CCG	TAT	GTT	ATT	GCC	GGG	AAA	1730
Tyr	Thr	Pro	Phe	Glu	Ala	Asp	Val	Thr	Pro	Tyr	Val	Ile	Ala	Gly	Lys	
340					345					350					355	
AGT	GTA	CGT	ATC	ACC	GTT	TGT	GTG	AAC	AAC	GAA	CTG	AAC	TGG	CAG	ACT	1778
Ser	Val	Arg	Ile	Thr	Val	Cys	Val	Asn	Asn	Glu	Leu	Asn	Trp	Gln	Thr	
			360					365						370		
ATC	CCG	CCG	GGA	ATG	GTG	ATT	ACC	GAC	GAA	AAC	GGC	AAG	AAA	AAG	CAG	1826
Ile	Pro	Pro	Gly	Met	Val	Ile	Thr	Asp	Glu	Asn	Gly	Lys	Lys	Lys	Gln	
			375					380					385			

TCT	TAC	TTC	CAT	AAT	TTC	TTT	AAC	TAT	GCC	GGG	ATC	CAT	CGC	AGC	GTA	1874
Ser	Tyr	Phe	His	Asn	Phe	Phe	Asn	Tyr	Ala	Gly	Ile	His	Arg	Ser	Val	
		390					395					400				
ATG	CTC	TAC	ACC	ACG	CCG	AAC	ACC	TGG	GTG	GAC	GAT	ATC	ACC	GTG	GTG	1922
Met	Leu	Tyr	Thr	Thr	Pro	Asn	Thr	Trp	Val	Asp	Asp	Ile	Thr	Val	Val	
	405					410					415					
ACG	CAT	GTC	GCG	CAA	GAC	TGT	AAC	CAC	GCG	TCT	GTT	GAC	TGG	CAG	GTG	1970
Thr	His	Val	Ala	Gln	Asp	Cys	Asn	His	Ala	Ser	Val	Asp	Trp	Gln	Val	
420					425					430					435	
GTG	GCC	AAT	GGT	GAT	GTC	AGC	GTT	GAA	CTG	CGT	GAT	GCG	GAT	CAA	CAG	2018
Val	Ala	Asn	Gly	Asp	Val	Ser	Val	Glu	Leu	Arg	Asp	Ala	Asp	Gln	Gln	
				440					445					450		
GTG	GTT	GCA	ACT	GGA	CAA	GGC	ACT	AGC	GGG	ACT	TTG	CAA	GTG	GTG	AAT	2066
Val	Val	Ala	Thr	Gly	Gln	Gly	Thr	Ser	Gly	Thr	Leu	Gln	Val	Val	Asn	
			455				460						465			
CCG	CAC	CTC	TGG	CAA	CCG	GGT	GAA	GGT	TAT	CTC	TAT	GAA	CTG	TGC	GTC	2114
Pro	His	Leu	Trp	Gln	Pro	Gly	Glu	Gly	Tyr	Leu	Tyr	Glu	Leu	Cys	Val	
		470					475						480			
ACA	GCC	AAA	AGC	CAG	ACA	GAG	TGT	GAT	ATC	TAC	CCG	CTT	CGC	GTC	GGC	2162
Thr	Ala	Lys	Ser	Gln	Thr	Glu	Cys	Asp	Ile	Tyr	Pro	Leu	Arg	Val	Gly	
	485					490					495					
ATC	CGG	TCA	GTG	GCA	GTG	AAG	GGC	GAA	CAG	TTC	CTG	ATT	AAC	CAC	AAA	2210
Ile	Arg	Ser	Val	Ala	Val	Lys	Gly	Glu	Gln	Phe	Leu	Ile	Asn	His	Lys	
500					505					510					515	
CCG	TTC	TAC	TTT	ACT	GGC	TTT	GGT	CGT	CAT	GAA	GAT	GCG	GAC	TTA	CGT	2258
Pro	Phe	Tyr	Phe	Thr	Gly	Phe	Gly	Arg	His	Glu	Asp	Ala	Asp	Leu	Arg	
				520					525					530		
GGC	AAA	GGA	TTC	GAT	AAC	GTG	CTG	ATG	GTG	CAC	GAC	CAC	GCA	TTA	ATG	2306
Gly	Lys	Gly	Phe	Asp	Asn	Val	Leu	Met	Val	His	Asp	His	Ala	Leu	Met	
			535					540					545			
GAC	TGG	ATT	GGG	GCC	AAC	TCC	TAC	CGT	ACC	TCG	CAT	TAC	CCT	TAC	GCT	2354
Asp	Trp	Ile	Gly	Ala	Asn	Ser	Tyr	Arg	Thr	Ser	His	Tyr	Pro	Tyr	Ala	
		550					555					560				
GAA	GAG	ATG	CTC	GAC	TGG	GCA	GAT	GAA	CAT	GGC	ATC	GTG	GTG	ATT	GAT	2402
Glu	Glu	Met	Leu	Asp	Trp	Ala	Asp	Glu	His	Gly	Ile	Val	Val	Ile	Asp	
	565					570					575					
GAA	ACT	GCT	GCT	GTC	GGC	TTT	AAC	CTC	TCT	TTA	GGC	ATT	GGT	TTC	GAA	2450
Glu	Thr	Ala	Ala	Val	Gly	Phe	Asn	Leu	Ser	Leu	Gly	Ile	Gly	Phe	Glu	
580					585					590					595	
GCG	GGC	AAC	AAG	CCG	AAA	GAA	CTG	TAC	AGC	GAA	GAG	GCA	GTC	AAC	GGG	2498
Ala	Gly	Asn	Lys	Pro	Lys	Glu	Leu	Tyr	Ser	Glu	Glu	Ala	Val	Asn	Gly	
				600					605					610		
GAA	ACT	CAG	CAA	GCG	CAC	TTA	CAG	GCG	ATT	AAA	GAG	CTG	ATA	GCG	CGT	2546
Glu	Thr	Gln	Gln	Ala	His	Leu	Gln	Ala	Ile	Lys	Glu	Leu	Ile	Ala	Arg	
			615					620					625			
GAC	AAA	AAC	CAC	CCA	AGC	GTG	GTG	ATG	TGG	AGT	ATT	GCC	AAC	GAA	CCG	2594
Asp	Lys	Asn	His	Pro	Ser	Val	Val	Met	Trp	Ser	Ile	Ala	Asn	Glu	Pro	
		630					635					640				
GAT	ACC	CGT	CCG	CAA	GGT	GCA	CGG	GAA	TAT	TTC	GCG	CCA	CTG	GCG	GAA	2642
Asp	Thr	Arg	Pro	Gln	Gly	Ala	Arg	Glu	Tyr	Phe	Ala	Pro	Leu	Ala	Glu	
	645					650					655					

GCA ACG CGT AAA CTC GAC CCG ACG CGT CCG ATC ACC TGC GTC AAT GTA	2690
Ala Thr Arg Lys Leu Asp Pro Thr Arg Pro Ile Thr Cys Val Asn Val	
660 665 670 675	
ATG TTC TGC GAC GCT CAC ACC GAT ACC ATC AGC GAT CTC TTT GAT GTG	2738
Met Phe Cys Asp Ala His Thr Asp Thr Ile Ser Asp Leu Phe Asp Val	
680 685 690	
CTG TGC CTG AAC CGT TAT TAC GGA TGG TAT GTC CAA AGC GGC GAT TTG	2786
Leu Cys Leu Asn Arg Tyr Tyr Gly Trp Tyr Val Gln Ser Gly Asp Leu	
695 700 705	
GAA ACG GCA GAG AAG GTA CTG GAA AAA GAA CTT CTG GCC TGG CAG GAG	2834
Glu Thr Ala Glu Lys Val Leu Glu Lys Glu Leu Leu Ala Trp Gln Glu	
710 715 720	
AAA CTG CAT CAG CCG ATT ATC ATC ACC GAA TAC GGC GTG GAT ACG TTA	2882
Lys Leu His Gln Pro Ile Ile Ile Thr Glu Tyr Gly Val Asp Thr Leu	
725 730 735	
GCC GGG CTG CAC TCA ATG TAC ACC GAC ATG TGG AGT GAA GAG TAT CAG	2930
Ala Gly Leu His Ser Met Tyr Thr Asp Met Trp Ser Glu Glu Tyr Gln	
740 745 750 755	
TGT GCA TGG CTG GAT ATG TAT CAC CGC GTC TTT GAT CGC GTC AGC GCC	2978
Cys Ala Trp Leu Asp Met Tyr His Arg Val Phe Asp Arg Val Ser Ala	
760 765 770	
GTC GTC GGT GAA CAG GTA TGG AAT TTC GCC GAT TTT GCG ACC TCG CAA	3026
Val Val Gly Glu Gln Val Trp Asn Phe Ala Asp Phe Ala Thr Ser Gln	
775 780 785	
GGC ATA TTG CGC GTT GGC GGT AAC AAG AAA GGG ATC TTC ACT CGC GAC	3074
Gly Ile Leu Arg Val Gly Gly Asn Lys Lys Gly Ile Phe Thr Arg Asp	
790 795 800	
CGC AAA CCG AAG TCG GCG GCT TTT CTG CTG CAA AAA CGC TGG ACT GGC	3122
Arg Lys Pro Lys Ser Ala Ala Phe Leu Leu Gln Lys Arg Trp Thr Gly	
805 810 815	
ATG AAC TTC GGT GAA AAA CCG CAG CAG GGA GGC AAA CAA TGAAGCTT	3169
Met Asn Phe Gly Glu Lys Pro Gln Gln Gly Gly Lys Gln	
820 825 830	

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val	
1 5 10 15	
Gly Asp Arg Val Thr Ile Thr Cys Ser Thr Ser Ser Ser Val Ser Tyr	
20 25 30	
Met His Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile	
35 40 45	
Tyr Ser Thr Ser Asn Leu Ala Ser Gly Val Pro Ser Arg Phe Ser Gly	
50 55 60	

Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro  
65 70 75 80  
Glu Asp Ile Ala Thr Tyr Tyr Cys His Gln Trp Ser Ser Tyr Pro Thr  
85 90 95  
Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro  
100 105 110  
Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr  
115 120 125  
Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys  
130 135 140  
Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu  
145 150 155 160  
Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser  
165 170 175  
Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala  
180 185 190  
Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe  
195 200 205  
Asn Arg Gly Glu Cys  
210

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 832 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Glu Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg Pro Ser  
1 5 10 15  
Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Thr Ile Ser Ser  
20 25 30  
Gly Tyr Ser Trp His Trp Val Arg Gln Pro Pro Gly Arg Gly Leu Glu  
35 40 45  
Trp Ile Gly Tyr Ile Gln Tyr Ser Gly Ile Thr Asn Tyr Asn Pro Ser  
50 55 60  
Leu Lys Ser Arg Val Thr Met Leu Val Asp Thr Ser Lys Asn Gln Phe  
65 70 75 80  
Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr  
85 90 95  
Cys Ala Arg Glu Asp Tyr Asp Tyr His Trp Tyr Phe Asp Val Trp Gly  
100 105 110  
Gln Gly Ser Leu Val Thr Val Thr Val Ser Ser Ala Ser Thr Lys Gly  
115 120 125  
Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg Ser Thr Ser Gly Gly  
130 135 140



Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	Glu	Pro	Val
145					150					155					160
Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	His	Thr	Phe
				165					170					175	
Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val
			180					185					190		
Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Thr	Cys	Asn	Val
		195					200					205			
Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Arg	Val	Glu	Leu	Gly
	210					215					220				
Ser	Gly	Ser	Gly	Ser	Met	Val	Arg	Pro	Val	Glu	Thr	Pro	Thr	Arg	Glu
225					230					235					240
Ile	Lys	Lys	Leu	Asp	Gly	Leu	Trp	Ala	Phe	Ser	Leu	Asp	Arg	Glu	Asn
				245					250					255	
Cys	Gly	Ile	Asp	Gln	Arg	Trp	Trp	Glu	Ser	Ala	Leu	Gln	Glu	Ser	Arg
			260					265					270		
Ala	Ile	Ala	Val	Pro	Gly	Ser	Phe	Asn	Asp	Gln	Phe	Ala	Asp	Ala	Asp
		275					280					285			
Ile	Arg	Asn	Tyr	Ala	Gly	Asn	Val	Trp	Tyr	Gln	Arg	Glu	Val	Phe	Ile
	290					295					300				
Pro	Lys	Gly	Trp	Ala	Gly	Gln	Arg	Ile	Val	Leu	Arg	Phe	Asp	Ala	Val
305					310					315					320
Thr	His	Tyr	Gly	Lys	Val	Trp	Val	Asn	Asn	Gln	Glu	Val	Met	Glu	His
				325					330					335	
Gln	Gly	Gly	Tyr	Thr	Pro	Phe	Glu	Ala	Asp	Val	Thr	Pro	Tyr	Val	Ile
			340					345					350		
Ala	Gly	Lys	Ser	Val	Arg	Ile	Thr	Val	Cys	Val	Asn	Asn	Glu	Leu	Asn
		355					360					365			
Trp	Gln	Thr	Ile	Pro	Pro	Gly	Met	Val	Ile	Thr	Asp	Glu	Asn	Gly	Lys
	370					375					380				
Lys	Lys	Gln	Ser	Tyr	Phe	His	Asn	Phe	Phe	Asn	Tyr	Ala	Gly	Ile	His
385					390					395					400
Arg	Ser	Val	Met	Leu	Tyr	Thr	Thr	Pro	Asn	Thr	Trp	Val	Asp	Asp	Ile
			405						410					415	
Thr	Val	Val	Thr	His	Val	Ala	Gln	Asp	Cys	Asn	His	Ala	Ser	Val	Asp
			420					425					430		
Trp	Gln	Val	Val	Ala	Asn	Gly	Asp	Val	Ser	Val	Glu	Leu	Arg	Asp	Ala
		435					440					445			
Asp	Gln	Gln	Val	Val	Ala	Thr	Gly	Gln	Gly	Thr	Ser	Gly	Thr	Leu	Gln
	450						455				460				
Val	Val	Asn	Pro	His	Leu	Trp	Gln	Pro	Gly	Glu	Gly	Tyr	Leu	Tyr	Glu
465					470					475					480
Leu	Cys	Val	Thr	Ala	Lys	Ser	Gln	Thr	Glu	Cys	Asp	Ile	Tyr	Pro	Leu
				485					490					495	

Arg	Val	Gly	Ile	Arg	Ser	Val	Ala	Val	Lys	Gly	Glu	Gln	Phe	Leu	Ile	500	505	510
Asn	His	Lys	Pro	Phe	Tyr	Phe	Thr	Gly	Phe	Gly	Arg	His	Glu	Asp	Ala	515	520	525
Asp	Leu	Arg	Gly	Lys	Gly	Phe	Asp	Asn	Val	Leu	Met	Val	His	Asp	His	530	535	540
Ala	Leu	Met	Asp	Trp	Ile	Gly	Ala	Asn	Ser	Tyr	Arg	Thr	Ser	His	Tyr	545	550	555
Pro	Tyr	Ala	Glu	Glu	Met	Leu	Asp	Trp	Ala	Asp	Glu	His	Gly	Ile	Val	565	570	575
Val	Ile	Asp	Glu	Thr	Ala	Ala	Val	Gly	Phe	Asn	Leu	Ser	Leu	Gly	Ile	580	585	590
Gly	Phe	Glu	Ala	Gly	Asn	Lys	Pro	Lys	Glu	Leu	Tyr	Ser	Glu	Glu	Ala	595	600	605
Val	Asn	Gly	Glu	Thr	Gln	Gln	Ala	His	Leu	Gln	Ala	Ile	Lys	Glu	Leu	610	615	620
Ile	Ala	Arg	Asp	Lys	Asn	His	Pro	Ser	Val	Val	Met	Trp	Ser	Ile	Ala	625	630	635
Asn	Glu	Pro	Asp	Thr	Arg	Pro	Gln	Gly	Ala	Arg	Glu	Tyr	Phe	Ala	Pro	645	650	655
Leu	Ala	Glu	Ala	Thr	Arg	Lys	Leu	Asp	Pro	Thr	Arg	Pro	Ile	Thr	Cys	660	665	670
Val	Asn	Val	Met	Phe	Cys	Asp	Ala	His	Thr	Asp	Thr	Ile	Ser	Asp	Leu	675	680	685
Phe	Asp	Val	Leu	Cys	Leu	Asn	Arg	Tyr	Tyr	Gly	Trp	Tyr	Val	Gln	Ser	690	695	700
Gly	Asp	Leu	Glu	Thr	Ala	Glu	Lys	Val	Leu	Glu	Lys	Glu	Leu	Leu	Ala	705	710	715
Trp	Gln	Glu	Lys	Leu	His	Gln	Pro	Ile	Ile	Ile	Thr	Glu	Tyr	Gly	Val	725	730	735
Asp	Thr	Leu	Ala	Gly	Leu	His	Ser	Met	Tyr	Thr	Asp	Met	Trp	Ser	Glu	740	745	750
Glu	Tyr	Gln	Cys	Ala	Trp	Leu	Asp	Met	Tyr	His	Arg	Val	Phe	Asp	Arg	755	760	765
Val	Ser	Ala	Val	Val	Gly	Glu	Gln	Val	Trp	Asn	Phe	Ala	Asp	Phe	Ala	770	775	780
Thr	Ser	Gln	Gly	Ile	Leu	Arg	Val	Gly	Gly	Asn	Lys	Lys	Gly	Ile	Phe	785	790	795
Thr	Arg	Asp	Arg	Lys	Pro	Lys	Ser	Ala	Ala	Phe	Leu	Leu	Gln	Lys	Arg	805	810	815
Trp	Thr	Gly	Met	Asn	Phe	Gly	Glu	Lys	Pro	Gln	Gln	Gly	Gly	Lys	Gln	820	825	830